

BlackBerry Gets Juiced

May 2007

At about 5 pm Pacific Daylight Time on Tuesday, April 17, 2007, BlackBerry messages stopped flowing across North America and elsewhere. Millions of BlackBerry subscribers were without service until Wednesday morning, when queued-up emails from the huge backlog of messages started to trickle through. It wasn't until Thursday that service was returned to normal.

How did this happen to a service that has come to be depended upon by such a large group of people?

The BlackBerry Success Story

BlackBerrys are handheld devices that are used to send and receive emails almost anywhere in the world. The BlackBerry service has been highly reliable and has attracted bankers, lawyers, journalists, law makers, and business people.



The BlackBerry devices and their associated services are provided by a Canadian company, Research in Motion Limited (RIM), of Waterloo, Ontario.

RIM is now in the process of going after the mainstream consumer market with its BlackBerry Pearl and BlackBerry 8800 devices, which add media players and mobile browsers to its email capability.

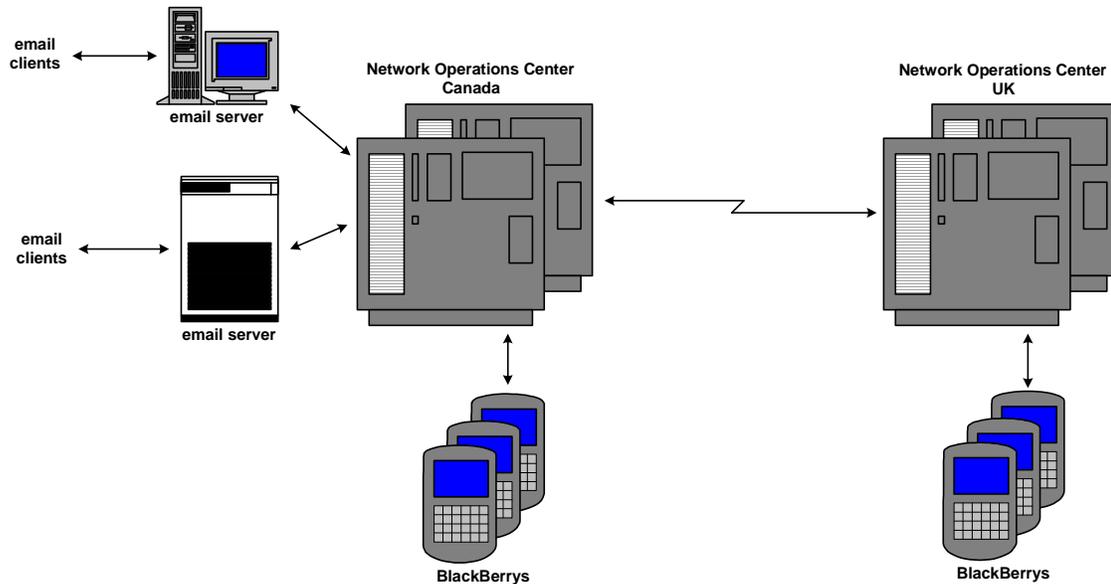
This marketing effort has been more than successful. From a subscriber base of just two million subscribers a little over a year ago, BlackBerry now boasts a quadrupling of that subscriber base to eight million subscribers; and it is currently signing up more than one million subscribers per quarter.

The BlackBerry Network

The BlackBerry network comprises two Network Operations Centers (NOC), one in Canada (Waterloo, Ontario) and the other in the UK. Each NOC acts as a hub, routing email messages between the BlackBerry handheld devices and corporate email servers and email accounts such as AOL and Yahoo.

The NOC in Canada serves BlackBerry subscribers in North America and parts of Asia. The NOC in the UK serves subscribers in Europe, Africa, and the Middle East.

BlackBerry Enterprise Servers receive emails from the various email subscribers and route encrypted emails to the appropriate NOC. The NOC authorizes users and forwards email messages to the appropriate BlackBerrys.



Likewise, messages received from the BlackBerrys are routed either to the destination email servers or to other BlackBerrys, as appropriate.

To accomplish this, each NOC interfaces with multiple mobile networks to get near-worldwide coverage.

The BlackBerry System Failure

On Tuesday, April 17, at 5 pm PDT, BlackBerry messages stopped flowing to subscribers in North America.¹ These subscribers could still make phone calls and send and receive emails, so it was clear that the mobile carriers were not at fault. However, they could not communicate over BlackBerry no matter which cellular carrier they used. Since it was North American BlackBerry subscribers that were affected, it was clear that the problem had to be in BlackBerry's Canadian NOC, which serves North America.

However, RIM management refused to acknowledge that they had a problem (though I suspect that the computer room was in mayhem).

It wasn't until twelve hours later, on Wednesday morning, that RIM management acknowledged that they had, in fact, had an outage. But they gave no reason for it and no prognosis of what to expect.

By Wednesday morning, the system had been returned to service; but it had to deal with a huge backlog of emails that had accumulated during the outage. These emails started to trickle out

¹ Material for this article was derived from c/net news.com, businessweek.com, news.yahoo.com, marketwatch.com, cbsnews.com, and msnbc.com over the period from April 18 to April 20, 2007, and from cnn.com, February 13, 2003.

Wednesday morning, but new emails that came in entered the end of the queue and were delayed for hours. It was not until Thursday that operations returned to normal.

Once operations normalized, RIM management began to release details of the outage. They reported that the outage was triggered by the "introduction of a new noncritical system routine" designed to optimize cache performance. They had not expected this change to affect the regular operations of BlackBerry.

However, despite previous testing, the new system routine set off an unexpected chain reaction. It triggered a series of interaction errors between the system's operational database and cache and resulted in a temporary outage until the backup system could be brought online.

After the RIM technicians isolated the database problem and tried unsuccessfully to fix it, they began the failover process to the backup system. But horrors of horrors, the failover attempt failed, despite failover having been previously tested.

Clearly, RIM's testing proved to be insufficient. RIM management states that it has now identified several aspects of its testing, monitoring, and recovery processes that it plans to improve as a result of this incident.

The BlackBerry Public Relations Failure

The system failure was just one of the failures during this outage. The other was a near lack of communication from RIM management to its subscribers. RIM's tight-lipped response angered many. Grumbles could be heard from the White House to the Canadian Parliament.

It took twelve hours to simply acknowledge that there had been an outage, and then there were no details forthcoming concerning the cause or what subscribers could expect. It was not until the system had been returned to service that RIM management issued some details on the cause of the outage.

This has evidently been the practice of RIM management over the years. Though the BlackBerry system is seen as a reliable system by its subscribers, it has had failures in the past. For instance, there was a failure in February of 2003. In June of 2005, there were two failures. During one of the 2005 failures, it took twelve hours before RIM confirmed that there was a problem and then only in a cryptic way aimed at the technically proficient.

This current outage has only added to RIM's problems. In April of 2007, RIM had to restate its earnings for the last three years because of accounting errors related to stock options. Last year, it had to pay 612.5 million dollars (US) to settle a long-running patent dispute.

Lessons Learned

Among other lessons, there are two primary lessons to be taken from this incident: test and communicate.

Test

RIM was clearly deficient in its test procedures. It seems inconceivable to me that one could decide that a change to the handling of database cache was unlikely to affect the operational system. This change should have been approached with the utmost care and made only after exhaustive testing.

The other testing deficiency was proper failover testing. However, this is a common problem in data centers. Failover testing can be very disruptive and risky. During the failover process, which could take hours, the system is unavailable to its users. Furthermore, often the system must be failed back to the primary system, thus denying service to users once again.

Perhaps even worse, failover testing can be risky. It is a complex process which gets little testing. Should the failover process fail during a failover test, service may be denied to users for even a longer period of time.

For these reasons, failover testing is often simply ignored, or failover is only partially tested. It is hoped that (1) failover will never be needed and (2) if it is, it will magically work.

To alleviate this risk, several companies have undertaken major efforts to reduce switchover time. Some have reported reducing such times from hours to minutes through rearchitecting the system, improving failover procedures, and automating the failover process with scripts. If failover can be reduced to a few minutes, the risk and aggravation to users is significantly reduced.²

Communications

Consumers have come to expect close communication with their utilities. Should power fail or telephone service be interrupted, these utilities perhaps put as much effort into keeping consumers informed either directly or through the media as they do in fixing the problem.

Clearly, this is not a philosophy embraced by RIM management. With any service, the service provider should learn to try to fix the problem while simultaneously communicating with the media and with their users – just as power and telephone utilities do.

Other Observations

A weakness of the BlackBerry service, which may also contribute to its strengths, is that half of all message traffic must pass through one node. Should that node fail, the service is out for those subscribers.

Other email services do not pass emails through a central server. Companies can set up their own email services to support a broad range of PDAs such as HP's iPaq and Palm's Treo.

Still, as active/backup pairs, the BlackBerry NOCs should have extreme availability. But so far, they can probably claim a little less than three 9s. Testing is one of their fundamental challenges to achieve better availability.

It is hard not to notice that RIM is operating two NOCs. If only these were architected to be an active/active pair, outages such as this could be eliminated. Software bugs and failover faults would be isolated to one node while all traffic was routed to the other node. Failover testing would be trivial. Extreme availability and subscriber satisfaction could then be truly achieved.

On a positive note, even though this was one of those famous "CNN Moments," it is interesting to note that RIM's stock price was not battered. It held fairly steady during the crisis and recovered quickly after an initial drop. All in all, this must be an indication of what people really think of BlackBerry. It may have its problems, but it is becoming one of those must-haves.

² See our article, "Tackling Switchover Times," in the October, 2006, issue of the Availability Digest.